Probabilidade e Estatística Matemática I

Parte 2: Variáveis Aleatórias, Funções de Distribuição e Esperanças

Silva, J.L.P.

Junho de 2022

Variáveis aleatórias e função de distribuição acumulada

Variável aleatória: definição

Definição para um dado espaço de probabilidade $(\Omega, \mathcal{A}, P)$, uma *variável aleatória* (*va*), denotada por X ou $X(\cdot)$ é uma função com domínio Ω e contradomínio a reta real. A função $X(\cdot)$ deve ser tal que o conjunto A_r, definido por $A_r = \{\omega : X(\omega) \leq r\}$, pertence a $\mathcal{A}$ para todo número real r.

- Uma *va* é, portanto, uma função do espaço amostral Ω nos reais, para a qual é possível calcular a probabilidade de ocorrência de seus valores.
- Apenas os elementos da σ-álgebra $\mathcal{A}$ têm atribuição de probabilidade.
- Em linguagem matemática mais formal, dizemos que uma v.a. é qualquer função real mensurável em $\mathcal{A}$.

Variável aleatória: exemplo

Exemplo Considere o experimento de lançar uma moeda. Seja a *va* X o número de caras. $\Omega = \{\text{cara, coroa}\}$, e $X(\omega) = 1$ se $\omega = \text{cara}$, e $X(\omega) = 0$, se $\omega = \text{coroa}$. Assim, a *va* X associa um número real a cada resultado do experimento.

Para verificar que X é uma *va* devemos mostrar que ela satisfaz a definição. Ou seja, devemos mostrar que $\{w : X(\omega) \leq r\}$ pertence a $\mathcal{A}$ para todo número real r. Seja $\mathcal{A} = \{\phi, \{\text{cara}\}, \{\text{coroa}\}, \Omega\}$.

- Se $r < 0$, $\{\omega : X(\omega) \leq r\} = \phi$.
- Se $0 \leq r < 1$, $\{\omega : X(\omega) \leq r\} = \{\text{coroa}\}$.
- Se $r \geq 1$, $\{\omega : X(\omega) \leq r\} = \Omega = \{\text{cara, coroa}\}$.

Logo, para todo r o conjunto $\{w : X(\omega) \leq r\}$ pertence a $\mathcal{A}$; portanto, $X(\cdot)$ é uma *va*.

Função de distribuição acumulada: definição

Definição A *função de distribuição acumulada* (*fda*) de uma v.a. X, denotada por $F_X(\cdot)$, é definida como a função com domínio na reta real e contradomínio no intervalo $[0, 1]$ que satisfaz
$F_X(x) = P(X \leq x) = P(\{\omega : X(\omega) \leq x\})$ para todo número real x.

- Uma *fda* é definida de forma única para cada *va* e pode ser usada para calcular probabilidades de eventos definidos em termos de sua correspondente *va*.
- Diferentes *va*'s podem ter a mesma função de distribuição.

Função de distribuição acumulada: exemplo

Exemplo Considere o experimento de lançar uma moeda. Assuma que a moeda é honesta. Seja X o número de caras. Então,

$$F_X(x) = \begin{cases} 0 & \text{se} \quad x < 0 \\ 1/2 & \text{se} \quad 0 \leq x < 1 \\ 1 & \text{se} \quad x \geq 1. \end{cases}$$

Ou $F_X(x) = \frac{1}{2} I_{[0,1)}(x) + I_{[1,\infty)}(x)$, em $I_A(\omega)$ é a *função indicadora* de A, que vale 1 se $\omega \in A$ e 0 se $\omega \notin A$.

Função de distribuição acumulada: exemplo

Exemplo Considere o experimento de lançar dois dados honestos. Seja X_k o valor da face superior no k-ésimo lançamento para $k = 1, 2$. Note que X_1 e X_2 são duas *va*'s diferentes com a mesma *fda*, que vale $F_{X_k}(x) = \sum_{i=1}^{5} \frac{i}{6} I_{[i,i+1)]}(x) + I_{[6,\infty)]}(x)$.

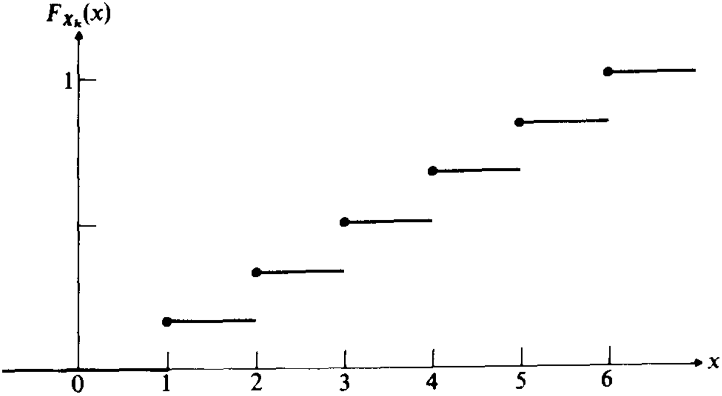


Figura 1: Fonte: Mood et. al (1974)

Função de distribuição acumulada: propriedades

1. $F_X(-\infty) \equiv \lim_{x \to -\infty} F_X(x) = 0$ e $F_X(+\infty) \equiv \lim_{x \to +\infty} F_X(x) = 1$.
2. $F_X(\cdot)$ é monótona não-decrescente; isto é, $F_X(a) \leq F_X(b)$ para $a < b$.
3. $F_X(\cdot)$ é contínua à direita; isto é,

$$\lim_{0 < h \to 0} F_X(x+h) = F_X(x).$$

Qualquer função $F(\cdot)$ com domínio na reta real e contradomínio no intervalo $[0, 1]$ satisfazendo as três propriedades acima é definida como uma *função de distribuição acumulada*.

Funções de densidade

Variáveis aleatórias discretas: definição

Definição Uma *va* X será definida como *discreta* se assume um número enumerável de valores (finito ou infinito).

- Assim, existe um número finito ou infinito enumerável de valores reais $x_1, x_2, x_3, \ldots$ que X pode assumir.
- Se X é discreta com valores distintos $x_1, x_2, \ldots, x_n, \ldots$, então $\Omega = \bigcup_n \{\omega : X(\omega) = x_n\} = \bigcup_n \{X = x_n\}$, e $\{X = x_i\} \cap \{X = x_j\} = \phi$ para $i \neq j$; logo $1 = P(\Omega) = \sum_n P(X = x_n)$ pelo terceiro axioma da probabilidade.

Função densidade discreta: definição

Definição Se X é discreta com valores $x_1, x_2, \ldots, x_n, \ldots$, então a função, denotada por $f_X(\cdot)$ e definida por

$$f_X(x) = \begin{cases} P(X = x_j) & \text{se} \quad x = x_j,\ j = 1, 2, \ldots, n, \ldots \\ 0 & \text{se} \quad x \neq x_j. \end{cases}$$

é chamada *função de densidade discreta* de X.

- Os valores da *va* discreta são chamados *pontos de massa* e $f_X(x_i)$ denota a *massa* associada com o ponto de massa x_i. A notação $p_X(\cdot)$ também é usada.
- $f_X(\cdot)$ é uma função com domínio na reta real e contradomínio $[0, 1]$.
- Podemos usar a notação $f_X(x) = \sum_{n=1}^{\infty} P(X = x_n) I_{\{x_n\}}(x)$, em que $I_{\{x_n\}}(x) = 1$ se $x = x_n$ e 0 se $x \neq x_n$.

Função densidade discreta: teorema

Teorema Seja X uma *va* discreta. $F_X(\cdot)$ pode ser obtida de $f_X(\cdot)$ e vice-versa.

Prova Denote os pontos de massa de X por $x_1, x_2, \ldots$

(ida) Suponha $f_X(\cdot)$ dada; então $F_X(x) = \sum_{\{j: x_j \leq x\}} f_X(x_j)$.

(volta) Por outro lado, suponha que $F_X(\cdot)$ é dada; então $f_X(x_j) = F_X(x_j) - \lim_{0<h\to 0} F_X(x_j - h)$; portanto, $f_X(x_j)$ pode ser encontrada para cada ponto de massa x_j. Como $f_X(x) = 0$ para $x \neq x_j,\ j = 1, 2, \ldots$, então $f_X(x)$ é determinada para todos os números reais.

Função densidade discreta: definição

Definição Qualquer função $f(\cdot)$ com domínio na reta real e contradomínio $[0,1]$ é definida como uma *função de densidade discreta* se, para algum conjunto enumerável, $x_1, x_2, \ldots, x_n, \ldots$

1. $f(x_j) > 0$ para $j = 1, 2, \ldots$
2. $f(x) = 0$ para $x \neq x_j,\ j = 1, 2, \ldots$
3. $\sum f(x_j) = 1$, em que o somatório se dá sobre os pontos $x_1, x_2, \ldots, x_n, \ldots$

- A definição permite falar de funções de densidade discretas sem alusão a variáveis aleatórias.
- Assim, podemos nos referir a propriedades que uma função de densidade deve ter sem nos referirmos a uma variável aleatória.

Variáveis aleatórias contínuas: definição

Definição A *va* X é chamada *contínua* se existe uma função $f_X(\cdot)$ tal que $F_X(x) = \int\limits_{-\infty}^{x} f_x(u)du$ para cada número real u. A função de distribuição acumulada $F_X(\cdot)$ de uma *va* contínua X é chamada *absolutamente contínua*.

Definição Se X é uma *va* contínua, a função $f_X(\cdot)$ em $F_X(x) = \int\limits_{-\infty}^{x} f_x(u)du$ é chamada de *função densidade de probabilidade* de X.

Variáveis aleatórias contínuas: teorema

Teorema Seja X uma *va* contínua. Então $F_X(\cdot)$ pode ser obtida de $f_X(\cdot)$ e vice-versa.

Prova *(ida)* Se X é uma *va* contínua e uma $f_X(\cdot)$ é dada, então $F_X(x)$ é obtida integrando $f_X(\cdot)$; isto é, $F_X(x) = \int_{-\infty}^{x} f_x(u)du$. *(volta)* Por outro lado, se $F_X(\cdot)$ é dada, então uma $f_X(x)$ pode ser obtida por derivação; isto é, $f_X(x) = dF_X(x)/dx$ para todos os pontos x em que $F_X(x)$ é diferenciável.

- Note que para uma *va* discreta, $f_X(x) = P(X = x)$.
- Para *va* contínua

$$f_X(x) = \frac{dF_X(x)}{dx} = \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x) - F_X(x - \Delta x)}{2\Delta x};$$

e, assim,$f_X(x)2\Delta x \approx F_X(x + \Delta x) - F_X(x - \Delta x) = P(x - \Delta x < X \leq \Delta x + \Delta x)$.

Variáveis aleatórias contínuas: exemplo

Exemplo Seja X a *va* que representa a duração de uma chamada telefônica. Podemos assumir que a distribuição e X é dada por $F_X(x) = (1 - e^{-\lambda x})I_{[0,\infty)}(x)$ para algum número real λ positivo. A função densidade de probabilidade correspondente é dada por $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$.

Se assumimos que a duração da chamada é medida em minutos, e se $\lambda = 1/5$,

- $P(5 < X \leq 10) = \int\limits_5^{10} \lambda e^{-\lambda x} dx = e^{-5\lambda} - e^{-10\lambda} = e^{-1} - e^{-2} =$ $0,368 - 0,135 = 0,233$ para $\lambda = 1/5$, ou
- $P(5 < X \leq 10) = P(X \leq 10) - P(X \leq 5) = (1 - e^{-10\lambda}) - (1 - e^{-5\lambda}) =$ $e^{-1} - e^{-2} = 0,233$ para $\lambda = 1/5$.

Variáveis aleatórias contínuas: definição

Definição Qualquer função $f(\cdot)$ com domínio na reta real e contradomínio $[0, \infty)$ é definida como uma *função densidade de probabilidade* se e somente se

(i) $f(x) \geq 0$ para todo x.

(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$.

- Com esta definição, podemos falar de função densidade de probabilidade sem referência a variáveis aleatórias.

Outros tipos de variáveis aleatórias

Nem todas as variáveis aleatórias são discretas ou contínuas, e nem todas as funções de distribuição são ou absolutamente contínuas ou discretas.

Vários exemplos práticos envolvem funções de distribuição acumuladas que são parcialmente discretas e parcialmente contínuas.

Há ainda funções de distribuição contínuas, chamadas de *singulares*, cuja derivada é 0 em quase todos os pontos[1]. A classificação como singular tem mais interesse teórico do que prático.

[1] "quase todos os pontos" significa que a propriedade só não é válida num conjunto de pontos com probabilidade nula.

Outros tipos de variáveis aleatórias: exemplo

Exemplo Considere o experimento de registrar o tempo que um motorista fica parado em determinada placa de PARE. Seja X a *va* que representa o tempo que decorre ao motorista fazer a parada.

- Há uma certa probabilidade de que não haverá tráfego de forma que o motorista prossiga sem precisar parar.
- Por outro lado, se o motorista tem que esperar, a espera pode assumir qualquer valor positivo.

O experimento pode ser modelado assumindo que X tem uma *fda* dada por $F_X(x) = (1 - pe^{-\lambda x})I_{[0,\infty)}(x)$. Esta $F_X(x)$ tem um salto de $1 - p$ em $x = 0$ mas é contínua para $x > 0$.

Outros tipos de variáveis aleatórias: exemplo

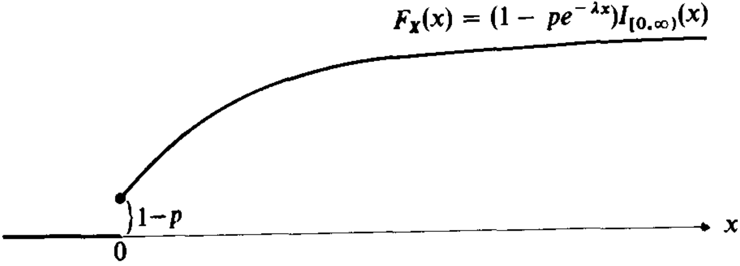


Figura 2: Fonte: Mood et. al (1974)

Decomposição de uma função de distribuição

Qualquer função de distribuição acumulada $F(x)$ pode ser representada na forma

$$F(x) = p_1 F^d(x) + p_2 F^{ac}(x) + p_3 F^{sc}(x), \quad p_i \geq 0 \ , i = 1, 2, 3,$$

com $\sum_{i=1}^{3} p_i = 1$. As funções $F^d(\cdot)$, $F^{ac}(\cdot)$, e $F^{sc}(\cdot)$ são as funções de distribuição discreta, absolutamente contínua e singular, respectivamente.

Decomposição de uma função de distribuição: exemplo

Exemplo No exemplo anterior, $F_X(x) = (1 - pe^{-\lambda x})I_{[0,\infty)}(x)$. Podemos decompor tal função de distribuição acumulada como

$$F_X(x) = (1-p)F^d(x) + pF^{ac}(x),$$

em que $F^d(x) = I_{[0,\infty)}(x)$ e $F^{ac}(x) = (1 - e^{-\lambda x})I_{[0,\infty)}(x)$.

Note que

$$\begin{aligned} F_X(x) =& (1-p)F^d(x) + pF^{ac}(x) \\ =& (1-p)I_{[0,\infty)}(x) + p(1-e^{-\lambda x})I_{[0,\infty)}(x) \\ =& (1-pe^{-\lambda x})I_{[0,\infty)}(x). \end{aligned}$$

Esperança e momentos

Média: definição

Definição Seja X uma *va*. A *média*, ou *esperança*, ou *valor esperado* de X, denotada por μ_X ou $E(X)$, é definida por

(i) Se X é discreta com pontos de massa $x_1, x_2, \dots, x_j, \dots$:

$$E(X) = \sum x_j f_X(x_j).$$

(ii) Se X é contínua com função densidade de probabilidade $f_X(x)$:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

(iii) Para uma *va* aleatória X arbitrária:

$$E(X) = \int_{0}^{\infty} [1 - F_X(x)] dx - \int_{-\infty}^{0} F_X(x) dx.$$

Média: exemplo

Exemplo Considere o experimento de lançar dois dados. Seja Y a diferença absoluta dos valores obtidos. A $f_Y(y)$ discreta é dada por

y	0	1	2	3	4	5
$f_Y(y)$	$6/36$	$10/36$	$8/36$	$6/36$	$4/36$	$2/36$

$$\begin{aligned} E(Y) = \sum y_j f_Y(y_j) = & \sum_{i=0}^{5} i f_Y(i) \\ = & 0 \times 6/36 + 1 \times 10/36 + \\ & 2 \times 8/36 + 3 \times 6/36 + \\ & 4 \times 4/36 + 5 \times 2/36 \\ = & 70/36 \end{aligned}$$

Média: exemplo

Exemplo Seja X uma *va* contínua com função de densidade $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$.

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{0}^{\infty} x\lambda e^{-\lambda x} dx = -xe^{-\lambda x}\Big|_0^{\infty} - \int_0^{\infty} (-e^{-\lambda x}) dx = 1/\lambda.$$

A função de distribuição acumulada é $F_X(x) = (1 - e^{-\lambda x}) I_{[0,\infty)}(x)$, logo

$$E(X) = \int_{0}^{\infty} [1 - F_X(x)] dx - \int_{-\infty}^{0} F_X(x) dx = \int_{0}^{\infty} (1 - 1 + e^{-\lambda x}) dx = 1/\lambda.$$

Média: exemplo

Exemplo Seja X uma *va* contínua com função de densidade $f_X(x) = x^{-2} I_{[1,\infty)}(x)$.

$$E(X) = \int_1^\infty x \frac{dx}{x^2} = \lim_{b\to\infty} \log b = \infty.$$

Dizemos que $E(X)$ não existe. Ou dizemos que a média de X é infinita já que a integral que define a média resulta em infinito.

Variância: definição

Definição Seja X uma *va* e seja μ_X sua média. A *variância* de X, denotada por σ_X^2 our $Var(X)$, é definida por:

1. Se X é discreta com pontos de massa $x_1, x_2, \ldots, x_j, \ldots$:

$$Var(X) = \sum_j (x_j - \mu_X)^2 f_X(x_j).$$

2. Se X é contínua com função densidade de probabilidade $f_X(x)$:

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx.$$

3. Para uma *va* aleatória X arbitrária:

$$Var(X) = \int_0^{\infty} 2x[1 - F_X(x) + F_X(-x)] dx - \mu_x^2.$$

Desvio padrão: definição

Definição Se X é uma *va*, o desvio padrão de X, denotado por σ_X é definido como $+\sqrt{Var(X)}$.

- O desvio padrão de uma *va*, assim como a variância, é uma medida de variabilidade ou dispersão dos seus valores.
- Em muitas aplicações, o desvio padrão é preferível à variância pois está na mesma unidade de medida da própria *va*.

Variância: exemplo

Exemplo Seja X o número total no lançamento de dois dados. Temos que $\mu_x = 7$ (confira!), logo

$$\begin{aligned}
Var(X) =& \sum (x - \mu_X)^2 f_x(x_j) \\
=& (2-7)^2 \times 1/36 + (3-7)^2 \times 2/36 + (4-7)^2 \times 3/36 + \\
& (5-7)^2 \times 4/36 + (6-7)^2 \times 5/36 + (7-7)^2 \times 6/36 + \\
& (8-7)^2 \times 5/36 + (9-7)^2 \times 4/36 + (10-7)^2 \times 3/36 + \\
& (11-7)^2 \times 2/36 + (12-7)^2 \times 1/36 \\
=& 210/36 \approx 5,833.
\end{aligned}$$

Média: exemplo

Exemplo Seja X uma *va* contínua com função de densidade $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$.

$$
\begin{aligned}
Var(X) &= \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \\
&= \int_{0}^{\infty} \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x} \\
&= 1/\lambda^2.
\end{aligned}
$$

Valor esperado de uma função de uma variável aleatória

Esperança de uma função de *va*: definição

Definição Seja X uma *va* e $g(\cdot)$ uma função com domínio e contradomínio na reta real. A *esperança* ou *valor esperado*, da função $g(\cdot)$ da *va* X, denotada por $E[g(X)]$ é definida por:

(i) Se X é discreta com pontos de massa $x_1, x_2, \ldots, x_j, \ldots$:

$$E[g(X)] = \sum_j g(x_j) f_X(x_j),$$

dado que a série seja absolutamente convergente.

(ii) Se X é contínua com função densidade de probabilidade $f_X(x)$:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx,$$

dado que $\int_{-\infty}^{\infty} |g(x)| f_X(x) dx < \infty$.

Propriedades do valor esperado: teorema

Teorema

1. $E[c] = c$, para uma constante c.
2. $E[cg(X)] = cE[g(X)]$, para uma constante c.
3. $E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E[g_1(X)] + c_2 E[g_2(X)]$.
4. $E[g_1(X)] \leq E[g_2(X)]$ se $g_1(X) \leq g_2(X)$.

Propriedades do valor esperado: demonstração

Prova Assuma X contínua.

(i) Tome $g(x) = c$, então

$$E[g(X)] = E[c] = \int_{-\infty}^{\infty} cf_X(x)dx = c\int_{-\infty}^{\infty} f_X(x)dx = c.$$

(ii)

$$E[cg(X)] = \int_{-\infty}^{\infty} cg(x)f_X(x)dx = c\int_{-\infty}^{\infty} g(x)f_X(x)dx = cE[g(X)].$$

Propriedades do valor esperado: demonstração

Prova (cont)

iii

$$
\begin{aligned}
E[c_1 g_1(X) + c_2 g_2(X)] &= \int_{-\infty}^{\infty} [c_1 g_1(x) + c_2 g_2(x)] f_X(x) dx \\
&= c_1 \int_{-\infty}^{\infty} g_1(x) f_X(x) dx + \int_{-\infty}^{\infty} c_2 g_2(x) f_X(x) dx \\
&= c_1 E[g_1(X)] + c_2 E[g_2(X)]
\end{aligned}
$$

iv Temos que $g_1(X) \leq g_2(X)$, logo

$$0 \leq E[g_2(X) - g_1(X)] = E[g_2(X)] - E[g_1(X)].$$

Provas similares podem ser feitas para o caso discreto.

Valor esperado: teorema

Teorema Se X é uma *va*, então
$Var(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2$, dado que $E(X^2)$ existista[2].

Prova Pela definição de variância e $E[g(X)]$, segue que $Var(X) = E[X - E(X)]^2$, e assim,

$$\begin{aligned} Var(X) = E[X - E(X)]^2 =& E[X^2 - 2XE(X) + (E(X))^2] \\ =& E(X^2) - 2[E(X)]^2 + (E(X))^2 \\ =& E(X^2) - [E(X)]^2. \end{aligned}$$

[2] Pode ser mostrado que se $E(X^2)$ existe então, $E(X)$ também existe.

Desigualdade: teorema

Teorema Seja X uma *va* e $g(\cdot)$ uma função não-negativa com domínio na reta real, então

$$P(g(X) \geq k) \leq \frac{E[g(X)]}{k}, \quad \text{para todo } k > 0.$$

Prova Assuma X contínua com densidade $f_X(\cdot)$, então

$$\begin{aligned}
E[g(X)] &= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\
&= \int_{x:g(x)<k} g(x) f_X(x) dx + \int_{x:g(x)\geq k} g(x) f_X(x) dx \\
&\geq \int_{x:g(x)\geq k} g(x) f_X(x) dx \\
&\geq \int_{x:g(x)\geq k} k f_X(x) dx = kP[g(X) \geq k].
\end{aligned}$$

Dividir por k completa a prova. A prova para o caso discreto é similar.

Desigualdade: corolário

Corolário Desigualdade de Chebyshev Seja X uma *va* com variância finita, então

$$P(|X-\mu_X| \geq r\sigma_X) = P\left[(X-\mu_X)^2 \geq r^2\sigma_X^2\right] \leq \frac{1}{r^2}, \quad \text{para todo } r > 0.$$

Prova Tome $g(x) = (x-\mu_X)^2$ e $k = r^2\sigma_x^2$ no teorema anterior.

Desigualdade de Chebyshev: comentários

A desigualdade de Chebyshev afirma que $P(|X - \mu_X| < r\sigma_X) \geq 1 - \frac{1}{r^2}$. Assim,

$$P(\mu_X - r\sigma_X < X < \mu_X + r\sigma_X) \geq 1 - \frac{1}{r^2}.$$

- Para $r = 2$, por exemplo, temos $P(\mu_X - 2\sigma_X < X < \mu_X + 2\sigma_X) \geq 3/4$; ou seja, para qualquer *va* X com variância finita pelo menos $3/4$ das observações devem estar dentro de dois devios padrões da sua média.

- Note que a desigualdade nos dá um limite, que não depende da distribuição de X, para a probabilidade de eventos descritos em termos da *va* e sua média e variância.

Desigualdade de Jensen

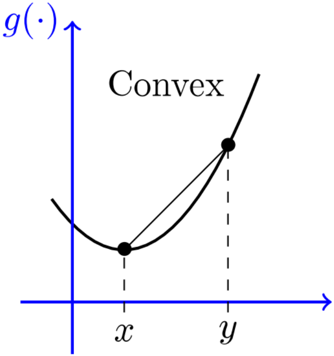
Definição Função convexa Uma função contínua $g(\cdot)$ com domínio e contradomínio na real real é chamda *convexa* se, para todo x_0 na reta real, existe uma linha que passa pelo ponto $(x_0, g(x_0))$ e está sob ou abaixo do gráfico da função $g(\cdot)$.

Teorema Desigualdade de Jensen Seja X uma *va* com média $E(X)$, e seja $g(\cdot)$ uma função convexa, então

$$E[g(X)] \geq g(E(X)).$$

Em geral, $E[g(X)] \neq g(E(X))$. Por exemplo, note que $g(x) = x^2$ é convexa; portanto a variância de X, que é $E(X^2) - [E(X)]^2$, é não negativa!

Desigualdade de Jensen: função convexa[3]



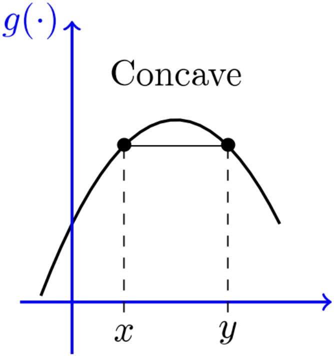


Figura 3: Representação de função côncava e convexa

[3] Extraído de
https://www.probabilitycourse.com/chapter6/6_2_5_jensen's_inequality.php

Momentos e função geradora de momentos

Momentos: definição

Definição Momento Se X é uma *va*, o *momento de ordem r* de X, geralmente denotado por $\mu_r^{'}$, é definido como

$$\mu_r^{'} = E(X^r),$$

se o momento existir. Note que $\mu_1^{'} = E(X) = \mu_X$.

Definição Momento central Se X é uma *va*, o *momento central de ordem r* de X, em torno de a é definido como $E((X-a)^r)$. Se $a = \mu_x$, temos o *momento central de ordem r* de X em torno de μ_x, denotado por μ_r, que é

$$\mu_r = E((X-\mu_x)^r).$$

- Note que $\mu_1 = E(X-\mu_x) = 0$, e $\mu_2 = E((X-\mu_x)^2) = Var(X)$.
- Todos os momentos ímpares de X em torno de μ_X são 0 se a densidade de X é simétrica em torno de μ_x, dado que tais momentos existam.

Quantis: definição

Definição O quantil q de uma *va* X ou de sua correspondente distribuição é denotado por ξ_q, e é definido como o menor ξ satisfazendo $F_X(\xi) \geq q$.

Se X é uma *va* contínua, o quantil q de X é dado pelo menor número ξ satisfazendo $F_X(\xi) = q$.

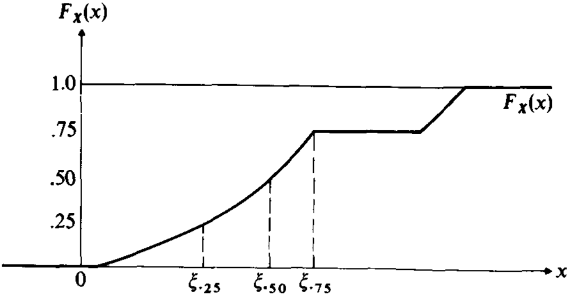


Figura 4: Fonte: Mood et. al (1974)

Mediana: definição

Definição A mediana de uma *va* X, denotada por med_X, $med(X)$, ou $\xi_{0,50}$ é o quantil $0,50$.

- Em alguns textos a mediana de X é definida como qualquer número, $med(X)$ satisfazendo

$$P(X \leq med(X)) \geq 1/2 \quad \text{e} \quad P(X \geq med(X)) \geq 1/2.$$

- Se X é uma *va* contínua, então a mediana de X satisfaz

$$\int_{-\infty}^{med(X)} f_X(x)dx = \frac{1}{2} = \int_{med(X)}^{-\infty} f_X(x)dx.$$

Função geradora de momentos: definição

Definição Seja X uma *va* com densidade $f_X(\cdot)$. O valor esperado de e^{tX} é definido como a *função geradora de momentos* de X se o valor esperado existe para t em algum intervalo $-h < t < h$, $h > 0$. A função geradora de momentos, denotada por $m_X(t)$ ou $m(t)$, é

$$m(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tX} f_X(x)dx,$$

se a *va* X é contínua e

$$m(t) = E(e^{tX}) = \sum_x e^{tX} f_X(x),$$

se a *va* X é discreta.

Função geradora de momentos

Se derivarmos a função geradora de momentos r vezes em relação a t, obtemos

$$\frac{d^r}{dt^r}m(t) = \int_{-\infty}^{\infty} x^r e^{tX} f_X(x)dx,$$

e fazendo $t \to 0$, encontramos

$$\frac{d^r}{dt^r}m(0) = E(X^r) = \mu_r^{'}.$$

Os momentos de uma distribuição podem ser obtidos da função geradora de momentos por diferenciação, o que justifica seu nome.

Função geradora de momentos

Se substituirmos e^{tX} por sua expansão em série, obtemos a expansão em série de $m(t)$ em termos dos momentos de $f_X(\cdot)$, isto é,

$$
\begin{aligned}
m(t) =& E\left[1 + Xt + \frac{1}{2!}(Xt)^2 + \frac{1}{3!}(Xt)^3 + \ldots\right] \\
=& 1 + \mu_1^{'} t + \frac{1}{2!}\mu_2^{'} t^2 + \ldots \\
=& \sum_{i=0}^{\infty} \frac{1}{i!}\mu_i^{'} t^i,
\end{aligned}
$$

do que fica evidente que $\mu_r^{'}$ pode ser obtida de $m(t)$.

Função geradora de momentos: exemplo

Exemplo Seja X uma *va* com densidade contínua $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$.

Temos

$$m_X(t) = E(e^{tX}) = \int_0^\infty e^{tX} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad \text{para} \quad t < \lambda.$$

Assim,

$$m^{'}(t) = \frac{dm_X(t)}{dt} = \frac{\lambda}{(\lambda - t)^2}, \quad \text{logo} \quad m^{'}(0) = E(X) = \frac{1}{\lambda}.$$

Função geradora de momentos: exemplo

Exemplo Seja X uma *va* com densidade discreta

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad \text{para} \quad x = 0, 1, 2, \ldots$$

Então

$$m_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)}.$$

Utilize a função geradora de momentos acima para encontrar $E(X)$, $E(X^2)$ e $Var(X)$.

Função geradora de momentos: teorema

Teorema Sejam X e Y são duas *va* com densidades $f_X(\cdot)$ e $f_Y(\cdot)$ respectivamente. Suponha que ambas $m_X(t)$ e $m_Y(t)$ existam e são iguais para todo t no intervalo $-h < t < h$, para algum $h > 0$. Então as duas funções de distribuição acumuladas $F_X(\cdot)$ e $F_Y(\cdot)$ são iguais.

- Se a função geradora de momentos de uma *va* existe, então esta função geradora de momentos determina de forma única a função de distribuição correspondente.
- Contudo, uma sequência de momentos $\mu_1^{'}, \mu_2^{'}, \ldots$ não determina de forma única uma função de distribuição.

Função geradora de momentos: exemplo

Exemplo Suponha que uma *va* X tenha função geradora de momentos $m_X(t) = 1/(1-t)$ para $-1 < t < 1$.

Sabemos que densidade de X é dada por $f_X(x) = e^{-x} I_{[0,\infty)}(x)$ pois vimos que $\lambda e^{-\lambda x} I_{[0,\infty)}(x)$ tem função geradora de momentos $\lambda/(\lambda - t)$.